Supplemental Information

Single-cell Transcriptomics Resolves Intermediate Glial Progenitors and Uncovers a Pivotal Determinant of Cell Fate and Gliomagenesis

Qinjie Weng, Jincheng Wang, Jiajia Wang, Danyang He, Zuolin Cheng, Feng Zhang, Ravinder Verma, Lingli Xu, Xinran Dong, Yunfei Liao, Xuelian He, Andrew Potter, Liguo Zhang, Chuntao Zhao, Mei Xin, Qian Zhou, Bruce J. Aronow, Perry J. Blackshear, Jeremy N. Rich, Qiaojun He, Wenhao Zhou, Mario L. Suvà, Ronald R. Waclaw, S. Steven Potter, Guoqiang Yu, and Q. Richard Lu

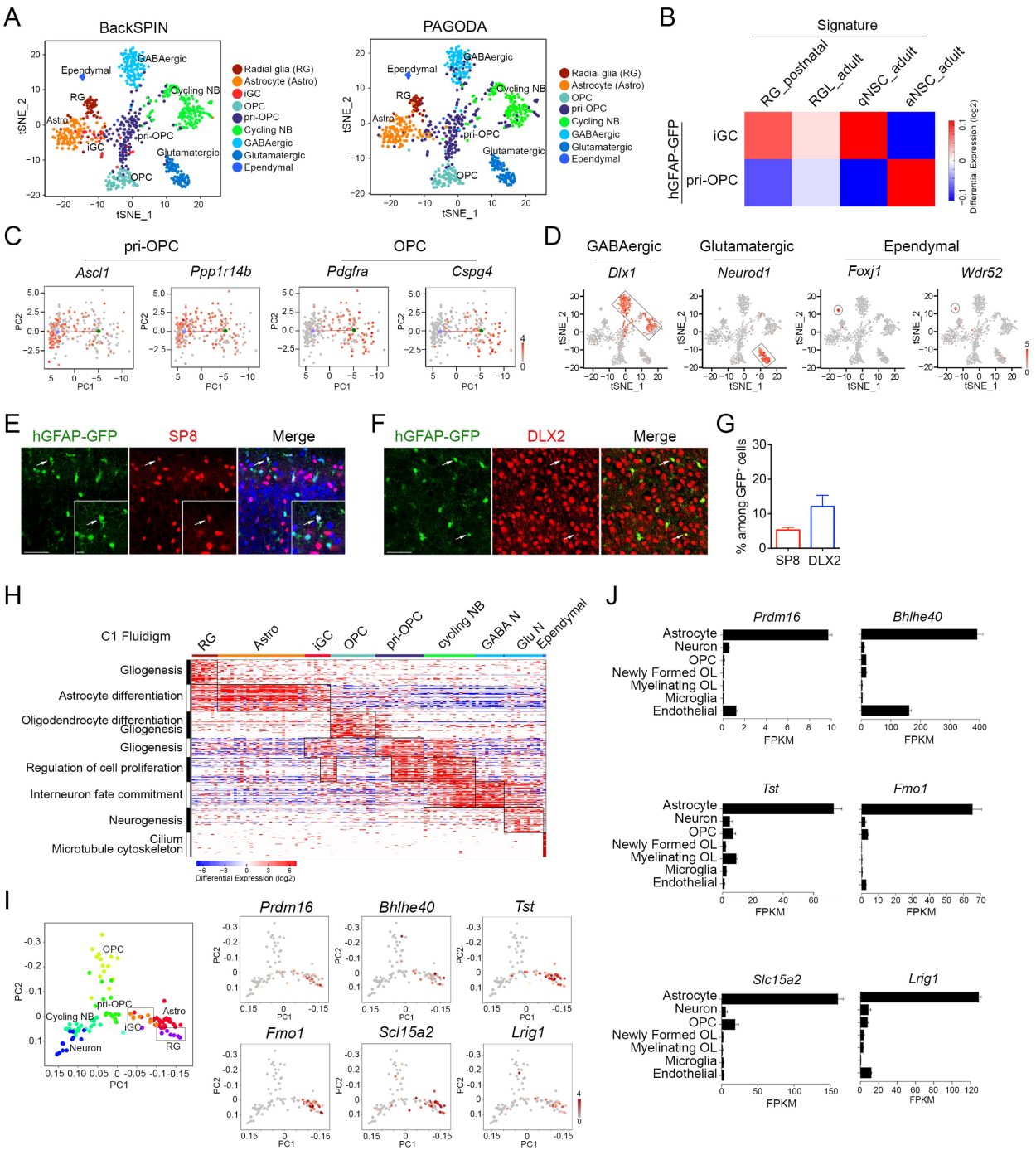

Figure S1. Identification of classification and potential astrocyte markers in hGFAP-GFP$^+$ cells. Related to Figure 1.

(A) Validation of Seurat analysis for single-cell transcriptomics of hGFAP-GFP$^+$ cells with different bioinformatics tools such as BackSPIN and PAGODA as indicated. The cell clusters are shown by tSNE. Different cell clusters are color coded and ordered as sequence from Seurat.

(B) Heatmap showing relative expression levels of the signature genes of postnatal RG and adult RG-like (RGL) cells (Hochgerner et al., 2018), as well as adult quiescent NSC (qNSC) and adult activated NSC (aNSC) cells (Dulken et al., 2017) in normal iGC and pri-OPC populations.

(C) Featureplots displaying the expression of selected marker genes for pri-OPC and OPC in the Slingshot predicted lineage trajectory.
(D) t-SNE plots showing expression of pan marker genes in neuron and ependymal clusters. The gene expression level is color-coded.
(E-F) Representative immunostained images showing expression of hGFAP-GFP, as well as SP8 (E) and Dlx2 (F) in cortical sections of P5 mice. Arrows indicate GFP^+ $Sp8^+$ or $DLX2^+$ co-labeling neuroblasts. Scale bars represent 50 and 10 μm, respectively.
(G) The percentage of $Sp8^+$ and $Dlx2^+$ neuronal cells among hGFAP-GFP^+ cells in P5 mouse cortices. Data are presented as mean ± SEM; n=3 animals.
(H) Heatmap of ICGS-ordered single-cell SMART-seq data from neonatal hGFAP-GFP^+ mice (n = 110 cells). Columns represent individual cells and rows represent genes. Gene-expression clusters were generated in AltAnalyze using HOPACH. ICGS-identified guide genes are indicated to the right. ICGS identified cell clusters are indicated above the plot. Gene ontology enrichment analysis for each HOPACH gene cluster is shown to the left.
(I) Left, PCA plot using single-value decomposition of all expressed genes in scRNA-seq libraries. Cells are colored according to their ICGS identities.
Right, visualization of transcription regulators that are highly enriched in astrocyte clusters. Each dot is a single cell, and cells are ordered by their relative position on PCA plot. Colored contours correspond to RNA abundance of indicated transcripts in each single cell.
(J) Gene expression of astrocyte marker in indicated cell types extracted from published database.

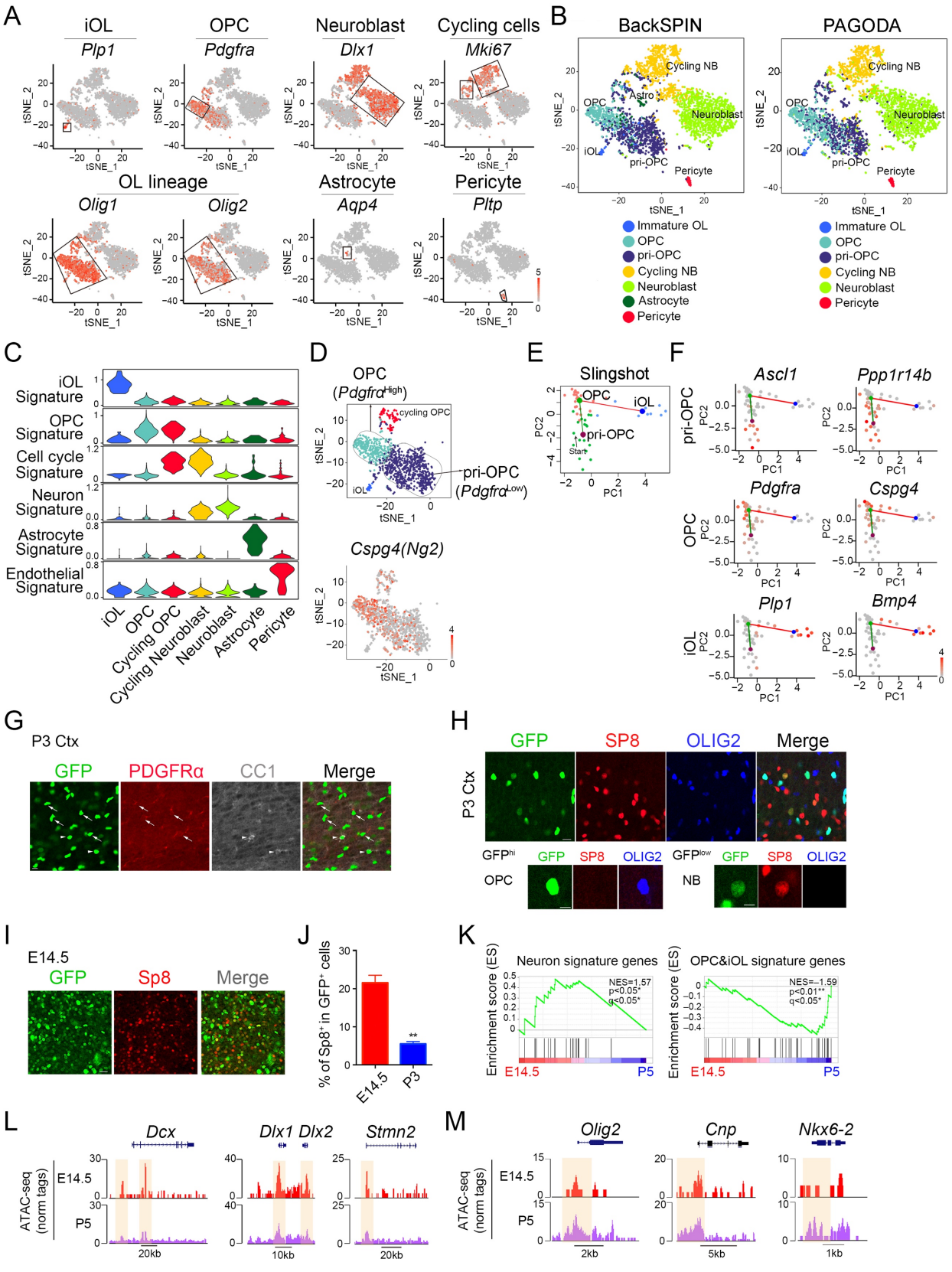
A
iOL
Plp1
OPC
Pdgfra
Neuroblast
Dlx1
Cycling cells
Mki67
OL lineage
Olig1
Olig2
Astrocyte
Aqp4
Pericyte
Pltp
tSNE_1
tSNE_2
B
BackSPIN
PAGODA
Cycling NB
Astro
OPC
Neuroblast
iOL
pri-OPC
Pericyte
Immature OL
OPC
pri-OPC
Cycling NB
Neuroblast
Astrocyte
Pericyte
C
iOL Signature
OPC Signature
Cell cycle Signature
Neuron Signature
Astrocyte Signature
Endothelial Signature
iOL
OPC
Cycling OPC
Cycling Neuroblast
Neuroblast
Astrocyte
Pericyte
D
OPC
(Pdgfrα^High)
cycling OPC
iOL
pri-OPC
(Pdgfrα^Low)
Cspg4(Ng2)
E
Slingshot
OPC
iOL
pri-OPC
Start
PC1
PC2
F
Ascl1
Ppp1r14b
Pdgfra
Cspg4
Plp1
Bmp4
pri-OPC
OPC
iOL
G
P3 Ctx
GFP
PDGFRα
CC1
Merge
H
GFP
SP8
OLIG2
Merge
P3 Ctx
GFP^hi
OPC
GFP^low
NB
I
E14.5
GFP
Sp8
Merge
J
% of Sp8+ in GFP+ cells
E14.5
P3
K
Neuron signature genes
OPC&iOL signature genes
Enrichment score (ES)
NES=1.57
p<0.05*
q<0.05*
NES=−1.59
p<0.01**
q<0.05*
E14.5
P5
L
Dcx
Dlx1 Dlx2
Stmn2
ATAC-seq (norm tags)
E14.5
P5
20kb
10kb
20kb
M
Olig2
Cnp
Nkx6-2
ATAC-seq (norm tags)
E14.5
P5
2kb
5kb
1kb

Figure S2. Validation of classification and gene identification of PDGFRα-GFP scRNA-seq dataset. Related to Figure 3.

(A) t-SNE plots showing expression of pan marker genes in distinct cell clusters. The gene expression level is color-coded.

(B) Validation of Seurat analysis for single-cell transcriptomics of PDGFRα-GFP$^+$ cells with different bioinformatics tools such as BackSPIN and PAGODA as indicated. The cell clusters are shown by tSNE. Different cell clusters are color-coded and ordered as sequence from Seurat.

(C) Violin plots showing the course of expression for established signature genes enriched in different cell types.

(D) t-SNE plot showing expression of *Cspg4* (Ng2) in OPC and pri-OPC cell clusters. The expression level is color coded.

(E) Slingshot predicts pseudo-time ordering of pri-OPC, OPC and iOL based on gene expression profiles in the PDGFRα-GFP$^+$ cells from Fluidigm C1 scRNA-seq dataset.

(F) Featureplots displaying the expression of selected marker genes for pri-OPC, OPC and iOL in the Slingshot predicted lineage.

(G) Representative images of immunolabeling of GFP (green), CC1 (white), and PDGFRα (red) in coronal brain sections from PDGFRα-GFP mice at P3. Scale bars represent 10 μm.

(H) Upper, Representative images of staining for Sp8 (red) and Olig2 (blue) on cortical sections from PDGFRα-GFP mice at P3. Nuclei are counterstained with DAPI (blue). Scale bars represent 10 μm. Lower, High-magnification images of GFP^{high}Olig2$^+$ OPCs and GFP^{low}Sp8$^+$ neuroblasts identified in PDGFRα-GFP mouse cortex at P3. Scale bars represent 5 μm.

(I) Representative images of cortical section from a PDGFRα-GFP mouse at E14.5 stained for Sp8 (red). Scale bars represent 10 μm.

(J) The percentage of Sp8$^+$ neuroblasts among PDGFRα-GFP$^+$ cells in E14.5 and P3 mouse cortex. Data are means ± SEM; n = 3 animals; ** $p<0.01$; two-tailed unpaired Student's t test.

(K) GSEA enrichment plots showing the comparison of ATAC-seq peak signaling of neuron and OPC&iOL related genes in E14.5 and P5 PDGFRα-GFP$^+$ single cell suspension as indicated. Data were generated from PDGFRa-GFP$^+$ cells of the cortices from a pool of 3 individual animals at each E14.5 and P5 stage, representing average peak signals.

(L) Genome browser tracks over indicated neuronal genes loci with ATAC-seq density (normalized tags) from E14.5 and P5 PDGFRα-GFP$^+$ cells.

(M) Genome browser tracks over OPC and iOL related genes loci with ATAC-seq density (normalized tags) at E14.5 and P5 PDGFRα-GFP$^+$ cells.

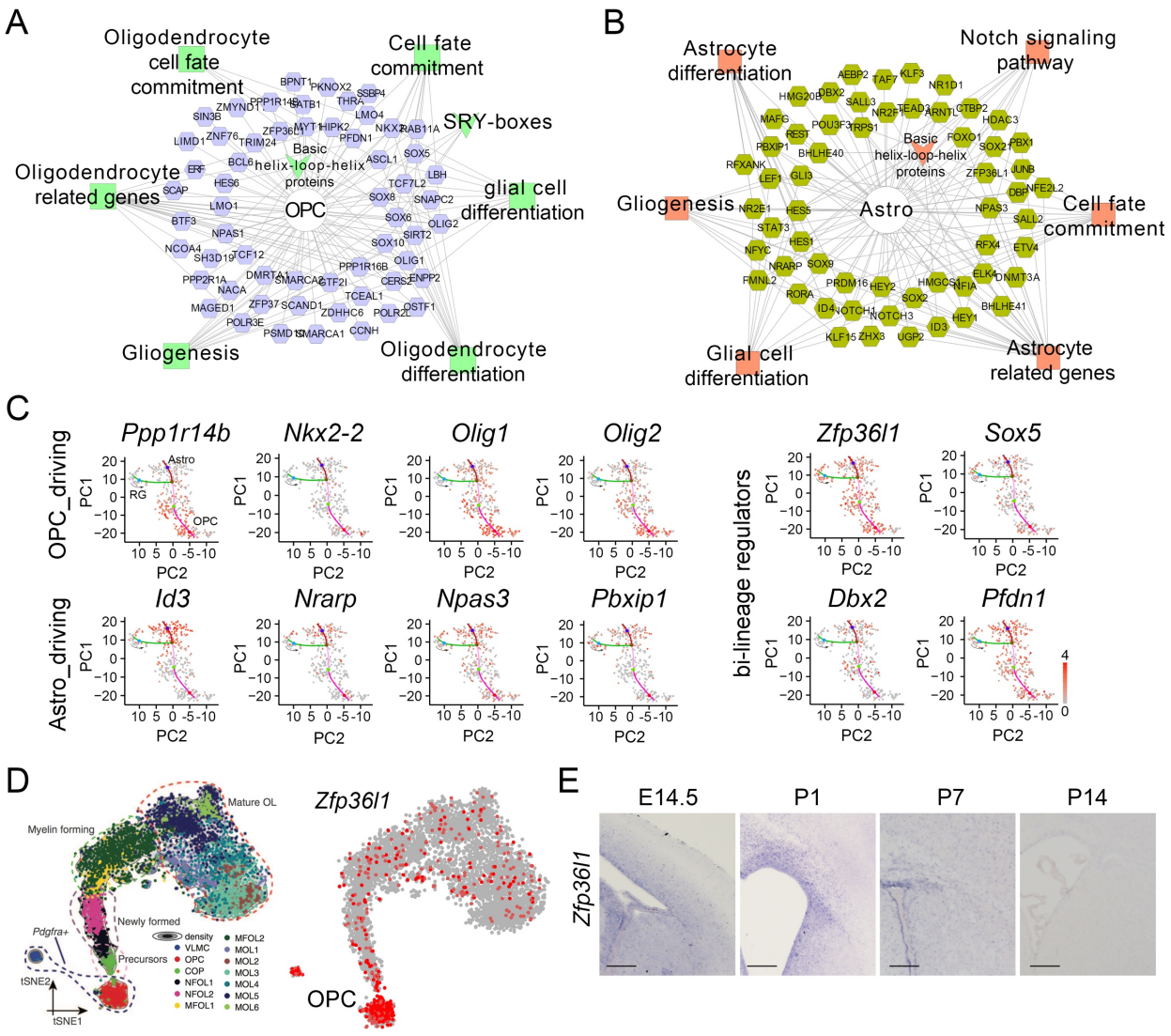

Figure S3. Transcriptional regulator network for driving OPC and astrocyte lineage. Related to Figure 4.

(A) ToppCluster analysis of representative OPC driving transcriptional regulators among the top Z-score gene list.

(B) ToppCluster analysis of representative Astro driving transcriptional regulators among the top Z-score gene list.

(C) Featureplots displaying the expression of selected lineage-driving regulators for OPC, astrocyte and bi-lineage regulators in the Slingshot predicted lineage trajectory.

(D) Left, distribution of distinct oligodendrocyte lineage cells based on the previous dataset (Marques et al., 2016). Right, t-SNE plot showing *Zfp36l1* gene expression enriched in OPC-like cells.

(E) In situ hybridization showing expression of *Zfp36l1* in the developing brain at the indicated stages. Scale bars, 200 μm.

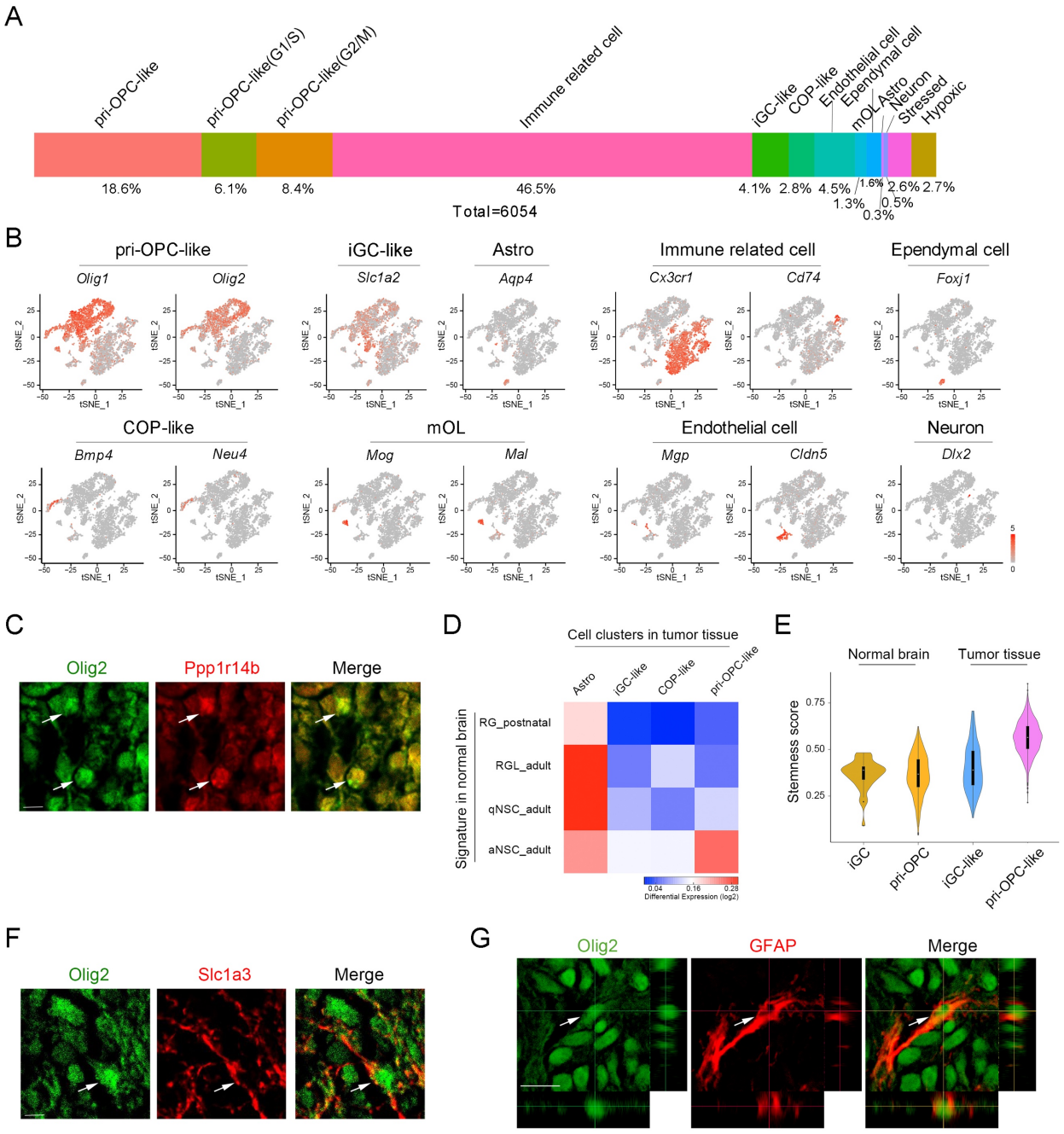

Figure S4. Glioblastoma cellular heterogeneity revealed by scRNA-seq at dpi 35. Related to Figure 5.

(A) Graph chart representing relative proportions of identified clusters in total single-cell pools in a mouse model of glioblastoma at dpi 35.

(B) t-SNE plots showing expression of marker genes in distinct cell clusters. The gene expression level is color coded.

(C) Representative images in murine glioma tissues showing expression of Ppp1r14b and Olig2. Arrows indicate the co-labeling cells. Scale bars, 5 μm.

(D) Heatmap showing the correlation of gene expression among various cell clusters in tumor tissues to the signature genes of postnatal RG and adult RG-like cells (Hochgerner et al., 2018), as well as adult qNSC and aNSC cells (Dulken et al., 2017) in the normal brain.
(E) Violin plot showing the distribution of the stemness score across the iGC and pri-OPC populations from normal developing brain and tumor tissues, respectively.
(F) Representative images in murine glioma tissues showing expression of Slc1a3/Olig2. Arrows indicate the co-labeling cells. Scale bar, 5 μm.
(G) Representative images in mouse glioma tissues showing expression of GFAP/Olig2. Arrows indicate the co-labeling cells. Orthogonal reconstruction of confocal images for a $GFAP^+$ /$Olig2^+$ cell by crosslines at the Z-axis level is shown. Scale bar, 10 μm.

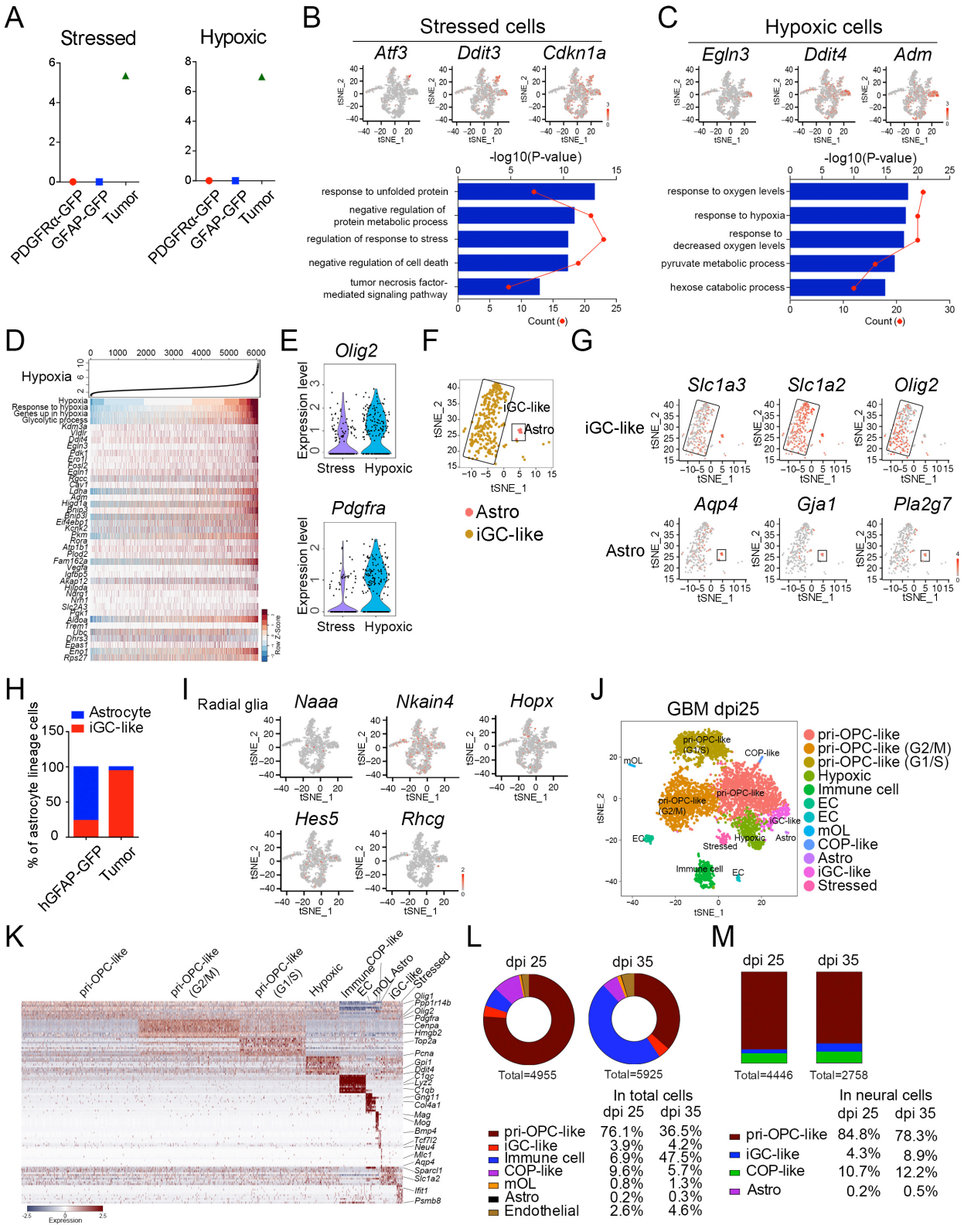

Figure S5. Identification of unique subpopulation of glioblastoma revealed by scRNA-seq. Related to Figure 6.

(A) Dot plot quantification of percentages of stressed cells and hypoxic cells in PDGFRα-GFP, hGFAP-GFP, and tumor cell dataset.
(B) Upper, t-SNE plots showing expression of stressed marker genes in tumor cells. The gene expression level is color-coded.
Lower, the gene ontology (GO) analysis of the top 50 variable genes in stressed cells.
(C) Upper, t-SNE plots showing expression of hypoxic marker genes in tumor cells. The gene expression level is color-coded.
Lower, the gene ontology (GO) analysis of the top 50 variable genes in hypoxia cells.
(D) Heatmap shows expression of the hypoxia signature, selected hypoxia gene sets and representative genes from the signature (rows) in individual glioblastoma cells (columns). Cells were ordered by hypoxia signature score.
(E) Violin plot showing the expression of the OPC markers *Olig2* and *PDGFRα* in hypoxic or stressed cells in tumor tissues.
(F) t-SNE plot based on all gene expression relationships in astrocyte and iGC-like cells from glioblastomas. Different cell clusters are labeled and color coded.
(G) t-SNE plot of representative marker genes in iGC-like and astrocyte cells from glioblastomas. The gene expression level in distinct cell clusters shown by t-SNE is color-coded.
(H) The ratio of iGC (or iGC-like) to Astro cells in hGFAP-GFP^{+} and tumor scRNA-seq pools.
(I) t-SNE plots showing selected marker genes of radial glia in tumor cells. The gene expression level is color-coded.
(J) t-SNE plot based on all gene expression relationships in single-cell data from glioblastomas at dpi 25. Different cell clusters are labeled and color coded. Tumor core tissues from 3 tumor-forming animals were collected.
(K) Heatmap of single-cell data from glioblastomas clustered using Seurat. Columns represent individual cells and rows represent genes. Gene-expression clusters are arranged based on most variable genes. Selected marker genes are displayed on the right.
(L) Pie chart representing cell percentage of distinct identified clusters in single-cell data of mouse glioma tissues at dpi 25 and dpi 35.
(M) The percentage of pri-OPC-like, iGC-like, COP-like, and astrocytes among neural cell groups in mouse glioma tissues at dpi 25 and dpi 35.

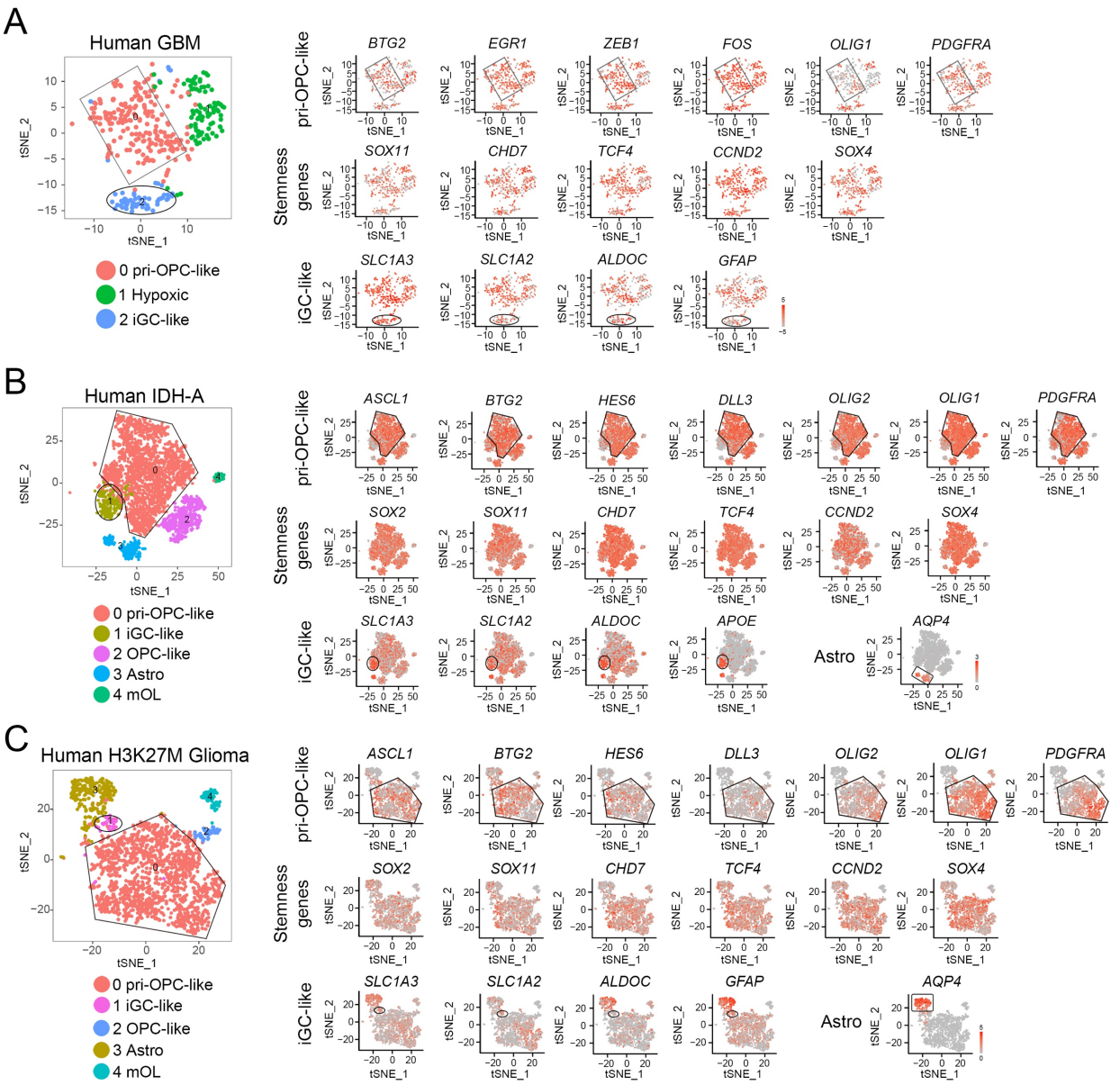

Figure S6. Meta-analysis of expression heterogeneity of human glioblastoma cells by scRNA-seq. Related to Figure 6.

(A) t-SNE plot based on unique signature gene expression relationships in single-cell data from human GBM cells. Different cell clusters are labeled and color coded, expression level of selected marker genes is color coded.

(B) t-SNE plot based on unique signature gene expression relationships in single-cell data from human IDH mutant astrocytoma cells. Different cell clusters are labeled and color coded, expression level of selected marker genes is color coded.

(C) t-SNE plot based on unique signature gene expression relationships in single-cell data from human H3K27M diffuse midline glioma cells. Different cell clusters are labeled and color coded, expression level of selected marker genes is color coded.

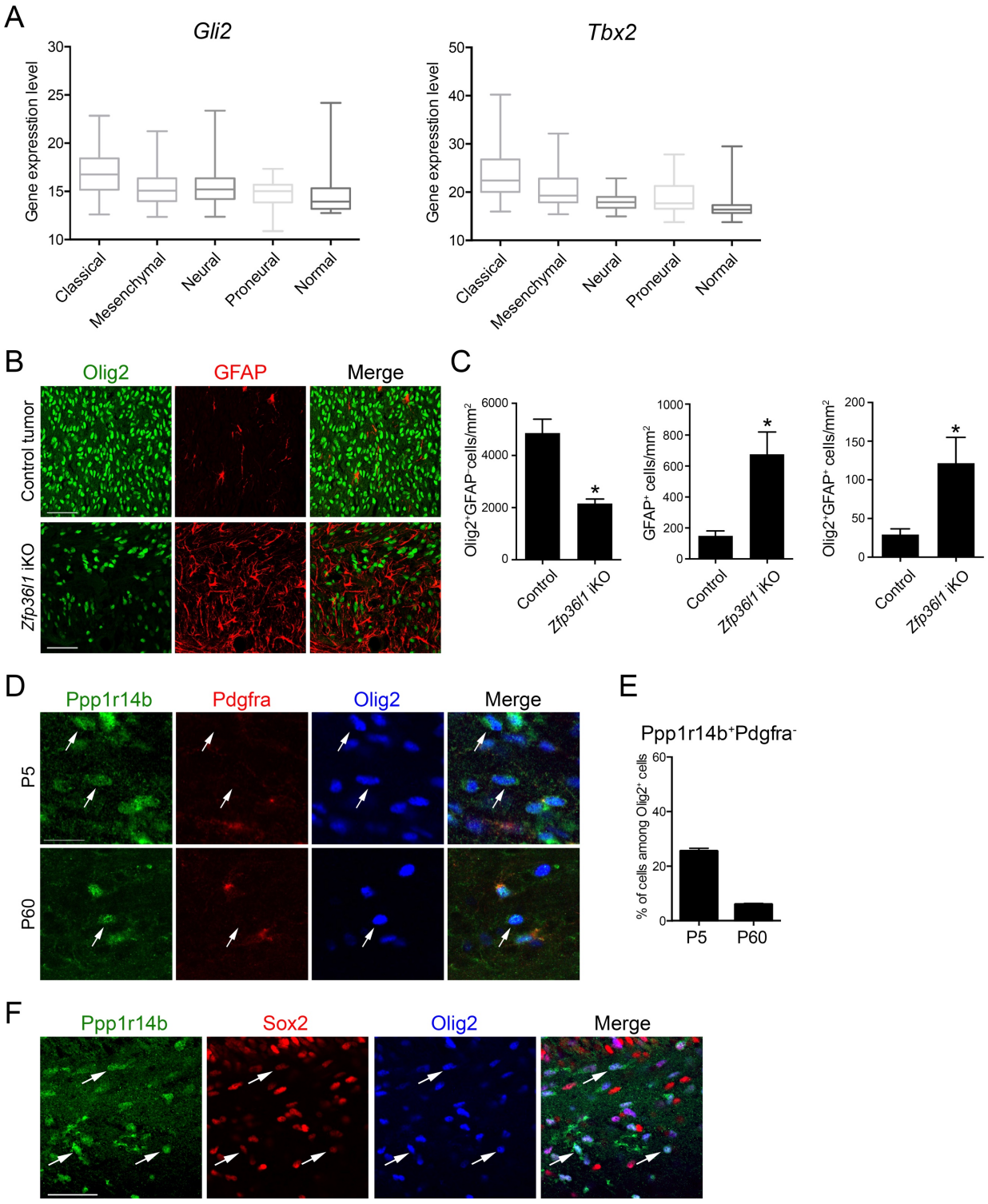

Figure S7. Characterization of Zfp36l1 function in tumor cell phenotype and Ppp1r14b-expressing cells. Related to Figure 7 and Discussion.
(A) Boxplot showing *Gli2* and *Tbx2* expression in distinct glioblastoma subtypes from TCGA dataset.

(B) Representative images of immunolabeling of Olig2 and GFAP in the tumor tissue from representative control tumor-forming mice at Dpi 40 and a *Zfp36l1* iKO mouse at Dpi 100. Scale bar represents 50 μm.
(C) Quantification of the number of $Olig2^{+}/GFAP^{-}$, $GFAP^{+}$ and $Olig2^{+}/GFAP^{+}$ cells per mm^2 within the tumor tissue of control-tumor forming mice (n = 3) at Dpi 40 and a *Zfp36l1* iKO mice (n = 2) at Dpi 100. * $p<0.05$; two-tailed unpaired Student's t test.
(D) Representative images showing expression of Ppp1r14b, Olig2 and Pdgfra in murine cortices at P5 and P60. Arrows indicate $Ppp1r14b^{+}Olig2^{+}Pdgfra^{-}$ pri-OPCs. Scale bar, 25 μm.
(E) The percentage of $Ppp1r14b^{+}Pdgfra^{-}$ cells among $Olig2^{+}$ cells. Data are presented as mean ± SEM; > 400 cell counts from 3 individual brains at each stage.
(F) Representative images showing co-expression of Ppp1r14b, Sox2 and Olig2 in the cortex from P5 mice. Arrows indicate the $Ppp1r14b^{+}Sox2^{+}Olig2^{+}$ cells. Scale bar, 50 μm.